DOCUMENT RESUME

ED 339 737 TM 017 656

AUTHOR Lawson, Stephen; And Others
TITLE Establishing the Psychometric Integrity of the Battelle Developmental Inventory for Young Children with Disabilities.
PUB DATE Nov 91
NOTE 47p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (20th, Lexington, KY, November 13, 1991) and presented in part at the Annual Meeting of the Allied Health Research Symposium (4th, Birmingham, AL, October 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Child Development; Construct Validity; *Diagnostic Tests; *Disabilities; Early Childhood Education; Early Intervention; Factor Analysis; Generalizability Theory; Longitudinal Studies; National Surveys; *Psychometrics; *Special Education; Test Norms; *Test Reliability; *Young Children
IDENTIFIERS *Battelle Developmental Inventory

ABSTRACT

Early childhood special educators recognize the necessity of establishing indices of reliability and validity for instruments that provide an index of developmental status. Many such instruments present little empirical evidence regarding psychometric integrity, particularly for a non-normative sample. The 341-item Battelle Developmental Inventory (BDI) is a commonly used instrument in early childhood special education research, but previously established indices of reliability and validity for this instrument were derived from samples of non-disabled children. This study investigated the item reliability and construct validity of the BDI for 78 children with severe disabilities. Data were collected as part of a longitudinal nationwide study investigating the efficacy of early intervention. Subjects were tested annually over a 5-year period with the BDI as one of several dependent measures, resulting in 265 cases across 8 age groups ranging from 11 months to 75 months. A generalizability study was conducted to establish the reliability of items; factor analytic procedures were used to determine construct validity. Results indicate that indices of reliability and validity differ from those obtained for the normative sample. Implications for use of the BDI and other developmental status measures in early childhood special education are discussed. (16 references; 14 data tables) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

BRUCE THOMPSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED339737

Establishing the Psychometric Integrity of the Battelle Developmental Inventory for Young Children with Disabilities

Stephen Lawson, Patricia Snyder, and Sarintha Stricklin
School of Allied Health Professions
Louisiana State University Medical Center
Bruce Thompson
Texas A&M University and Baylor University College of Medicine

Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Lexington, KY, November 13, 1991

TM017656

BEST COPY AVAILABLE

Abstract

Early childhood special educators recognize the necessity of establishing indices of reliability and validity for instruments that provide an index of developmental status. Many instruments used in early childhood special education present little empirical evidence regarding psychometric integrity, particularly for a non-normative sample. The Battelle Developmental Inventory (BDI) is a commonly used instrument in early childhood special education research. Yet, previously established indices of reliability and validity for this instrument were derived from samples of nondisabled children.

The present study investigated the item reliability and construct validity of the BDI for a sample of 78 children with disabilities. Data were collected as part of a longitudinal, nationwide study investigating the efficacy of early intervention. The authors conducted a generalizability study to establish the reliability of items; factor analytic procedures were used to determine construct validity. Results indicated that indices of reliability and validity differed from those obtained for the normative sample. Implications for use of the BDI and other developmental status measures in the field of early childhood special education efficacy research are discussed.

Establishing the Psychometric Integrity of the Battelle Developmental Inventory for Young Children with Disabilities

Establishment of measurement integrity is a critical component of all quantitative research. Without establishing the degree of psychometric integrity of instruments that are used to generate data, little confidence can be placed in the worth of the data (Kerlinger, 1986). Accurate conclusions concerning latent variables under consideration in a study cannot be made without relevant reliability and validity information for the measures used to quantify these variables.

The importance of this observation cannot be overemphasized. Sexton, Thompson, Scott, and Wood (1990) state that "These two psychometric properties (reliability and validity) are critical...because incorrect conclusions will be extrapolated from studies employing inadequate measures" (p. 37). Similarly, Kerlinger (1986) noted, "Poor measurement can invalidate any scientific investigation. There is growing understanding that all measuring instruments must be critically and empirically examined for their reliability and validity" (p. 431).

In the field of early childhood special education, the establishment of measurement integrity becomes even more relevant. Sheehan (1982) described five new directions in which professionals were moving in the area of infant assessment. One of these directions was the conduct of reliability and validity

studies on existing measures, using infants with disabilities, as opposed to the development of new instruments.

Researchers often use reliability and validity data reported elsewhere, such as in test manuals and in previous research efforts, when reporting the results of their investigation. However, **characteristics of measurement integrity relate to the data from the sample on which they were obtained, not to the test itself**. As Sax (1980) stated:

> It is more accurate to talk about the reliability of measurements (data, scores, and observations) than the reliability of tests (questions, items, and other tasks). Tests cannot be stable or unstable, but observations can. Any reference to the "reliability of a test" should always be interpreted to mean the "reliability of measurements or observations [i.e., a particular set of data] derived from a test". (p. 261).

Brennan (in press) concurred, noting that characteristics of reliability relate to the entire measurement procedure, not to the instrument.

High reliability of data, however, is no guarantee of valid scientific results (Kerlinger, 1986). Even data from a sample that can be considered reliable may not necessarily be valid. For example, a researcher may say that "the Smith Achievement Test (SAT) is reliable," when a variety of studies have shown that the test yields reliable data with diverse samples of high school seniors over repeated administrations under specified

conditions. A researcher would not presume that the same "reliable" test would yield meaningful data with first-graders.

Researchers in a variety of disciplines recognize that characteristics of reliability and validity relate to data obtained from normative or study samples, not to the test itself (cf. Payton, 1988; Sax, 1980). However, it is not uncommon to find references in the early childhood special education literature which belie this recognition. Researchers might state "the Smith Manual Dexterity Test is reliable and valid". Strictly speaking, this statement is imprecise. There is no harm in this practice as long as **our behavior** as early childhood special education researchers does not imply that we believe that tests themselves can be reliable and valid.

Reasons for Examining Instrument Integrity

Study samples used in early childhood special education frequently differ from normative samples. Normative samples are often drawn from a more global population of interest through a stratified or random selection process. Study samples in special education often are samples of convenience from a particular targeted population. Indices of reliability and validity that relate to the data obtained from the normative sample may, or may not, replicate with the study sample.

Researchers in this situation must therefore ask themselves, "How similar, or different, are the characteristics of my study sample from the normative sample?" "How might these similarities, or differences, impact indices of reliability and

validity?" "How well do results from past measurement studies generalize to current samples?" Researchers must realize that they are faced with two options: (a) to accept, solely on the basis of assumption, previously established indices of reliability and validity, or (b) to conduct a priori analyses that empirically evaluate the measurement integrity of the instrument for data obtained from the study sample.

It is our contention that the poorer the fit between the characteristics of the normative sample and the study sample, the more crucial these a priori analyses become. The reliability coefficients for data establish an upper bound on the effect sizes that can be detected in a study (Locke, Spirduso, & Silverman, 1987). Thus, reliability coefficients for the data obtained on study instruments used in the empirical investigation provide a basis for determining, a priori, whether substantive analyses are even plausible. These coefficients also allow the researcher to retrospectively interpret effect sizes (e.g., r^2) against the ceiling created by reliability coefficients.

In the remainder of the paper we describe two measurement integrity analyses using data obtained from a sample of children with severe disabilities on the Battelle Developmental Inventory (Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984). We discuss our rationale for conducting these analyses. We present the procedures used to conduct internal reliability and construct validity analyses for data obtained from the study sample.

Method

Instrumentation

The Battelle Developmental Inventory (BDI) is a standardized, individually administered assessment battery designed to measure key developmental skills in young children from birth through eight years of age. Early childhood special educators use the BDI for two primary purposes: (a) to establish eligibility for early intervention services based on comparisons between a specific child's performance and the normative sample, and (b) to determine the efficacy of various interventions used with young children with disabilities.

The BDI contains 341 items across five developmental domains: (a) Personal-social, (b) Adaptive, (c) Motor, (d) Communication, and (e) Cognitive. These five domains are further divided into 22 separate sub-domains. (See Appendix A for a complete listing of these sub-domains.) Basals and ceilings are established for a subject on each of the 22 sub-domains. Standard scores, percentile ranks, and age-equivalents can be calculated based on a subject's raw score in each sub-domain and domain, and for the total raw score across all domains.

The developers of the BDI provide estimates of inter-rater reliability (range = .70 - 1.0), test-retest reliability (range = .71 - .99), and standard error of measurement (range = 0.0 - 3.31) based on data obtained from the normative sample. Measures of validity, reported in the BDI manual for the same sample, include (a) content validity, verified by external content

experts; (b) construct validity, five factors including personal-social, adaptive, motor, and two communication factors; and (c) concurrent validity with the Vineland Social Maturity Scale subscales (Doll, 1965), ranging from .79 to .94.

Several researchers reported additional data regarding the concurrent validity of the BDI with other well-known early childhood measures for study samples containing children with disabilities (cf. Boyd, Welge, Sexton, & Miller, 1989; Sexton, Thompson, Perez, & Rheams, 1990). Correlations between subdomains on the BDI and these other measures are moderate to high and positive, indicating that the BDI has adequate criterion validity for selected non-normative samples.

Sample

The normative sample for the BDI consisted of a stratified sample of 800 children without disabilities. Children ranged in age from birth through 95 months, distributed in approximately equal numbers among 10 age groups. The normative sample controlled for sex and minority status within each age group and residence (urban-rural). In Table 1 we present descriptive information regarding the BDI normative sample.

Insert Table 1 about here

The sample for our study consisted of 78 children with severe disabilities. All of these children resided in an urban area. These children were part of a national, longitudinal

investigation examining the efficacy of early intervention. They were tested annually over a five year timespan with the BDI as one of several dependent measures. A variety of factors, such as unavailability for testing, unusable test protocols, etc. resulted in different numbers of BDI test administrations per year. This resulted in 265 cases across the five years that comprised the present study sample.

Taken together, all 265 cases were distributed in unequal numbers across eight age groups ranging from 11 months to 75 months. The majority of the cases in this sample consisted of children of minority groups (n = 186), with a slightly higher percentage of males than females. In Table 2 we present descriptive information regarding our study sample.

Insert Table 2 about here

Procedures

Differences between the normative sample and our sample warranted the investigation of measurement integrity prior to conducting any substantive analyses regarding the efficacy of early intervention using the BDI as a dependent measure. In the sections that follow, we describe the measurement integrity analyses performed to evaluate the internal consistency reliability and the construct validity of the BDI.

Construct validity analysis. We conducted two exploratory factor analyses using data obtained from our study sample to

determine if the underlying structure of the BDI was consistent with that reported by the test developers. Sexton et al. (1990, p. 45) noted that "factor structure results offer important evidence regarding the integrity of measures". So crucial is the use of this procedure that factorial validity has been referred to as a type of construct validity (Nunnally, 1972).

We used 231 items of the BDI across 265 cases in the factor analysis. All children participating in our study functioned at a developmental age of 60 months or less. We therefore deleted those items from our analysis that the developers of the BDI placed above the 60 month age range on the test protocol. We summed each subject's score on each of the 22 BDI sub-domains and used these 22 scores to conduct the analyses.

Both factor analytic procedures consisted of a principal components analysis, with results rotated to the varimax criterion. In the first analysis, each of the 265 cases were treated as if they were independent. In the second analysis, we used the SPSSX DISCRIMINANT procedure to obtain a pooled intercorrelation matrix among subscales, with each year's test administration serving as the grouping variable. A factor analysis of this pooled matrix was then calculated. Both factor analyses were performed using the SPSSX FACTOR procedure.

Reliability analysis. We used generalizability theory, in contrast to classical reliability theory, to determine the internal consistency of the items on the BDI. Generalizability theory has been described as an essential aspect of reliability

assessment (Thompson, 1989a; in press-a) and as such provides a method for most accurately estimating the reliability of instruments. Classical theory provides several estimates of measurement error: (a) internal reliability, (b) stability reliability, (c) inter-rater reliability and (d) equivalence reliability. Researchers derive these estimates independently. As a result, four completely different reliability estimates may be derived from the same set of scores on the same instrument, and results may be contradictory.

Generalizability theory simultaneously estimates the influence of each of these sources of error on the data set through the use of analysis of variance (ANOVA) logic. As explained by Shavelson, Webb, and Rowley (1989, p. 923):

> The extension of CT (Classical Theory) to GT (Generalizability Theory) acknowledges multiple influences on measurement variance. Whereas CT partitions variance into only two sources (true score variance and error variance), GT partitions score variance into many sources, corresponding to systematic variance among the objects of measurement, to multiple error sources, and to their interactions.

The power of Generalizability Theory lies in providing the researcher with the ability to estimate multiple sources of measurement error (i.e., both separately and in interaction with one another) and to determine the best measurement design in

which reliable scores can be obtained in the most efficient manner.

The GENOVA computer program, developed by Crick and Brennan (1983), yielded estimates of internal consistency reliability on the BDI for our study sample. We used the same number of items (v = 231) and subjects (n = 265) for this analysis as described in the section on construct validity analysis. We conducted separate generalizability analyses to determine the internal consistency reliability of items corresponding to the obtained factor solution.

Results

Tables 3 contains the matrix of inter-correlations among BDI subdomains as reported by the test developers. This table is organized into the five domains (i.e., Personal-Social, Adaptive, Motor, Communication and Cognitive) as originally reported by Newborg et al. (1984). The developers of the BDI state that each domain can be administered independently of the others. This statement assumes that these five domains are uncorrelated with one other.

Insert Table 3 about here

In Tables 4 and 5, we provide the inter-correlation matrices used to derive the factor solutions for our study sample. Table 4 presents the pooled correlation matrix. Table 5 presents the correlation matrix with each case treated independently across

all five years. Overall, inter-correlations among subdomains tended to be lower with the sample data as compared to the BDI normative data. Interested readers are encouraged to use these matrices to replicate the results described later in this paper.

Insert Tables 4 and 5 about here

Tables 6 and 7 contain the pre- and post-rotated structure matrices of the first factor analysis in which each of the 265 cases was treated as being independent. This analysis yielded a three factor solution explaining 82.7% of the original variance. Tables 8 and 9 contain the results of the second factor analysis, in which we used the pooled intercorrelation matrix. This analysis also yielded a three factor solution with 78.5% of the original variance explained.

Insert Tables 6 through 9 about here

To determine the degree of similarity between these two obtained factor structure matrices, we rotated the Table 7 results to the "best fit" position with results presented in Table 9. The cosines of the angles of these factors are correlation coefficients (Thompson, in press-b). Results indicated an high positive correlation between the three factors in the analyses, with obtained correlation coefficients of .9987, .9985, and .9974, respectively. In interpreting these obtained

factor solutions, Factors I and II appear to represent social-interactive, attentional, and problem solving skills. Factor III represents a distinct dimension related to motor proficiency and the execution of self-care activities.

Both of the results in the first two factor analyses differed from the five factor solution reported by the developers of the BDI. We conducted a final principal components analysis of the subscale intercorrelation matrix, as given in the BDI manual, to obtain a factor structure matrix for comparison purposes. However, results of this factor analysis yielded only a two factor solution. Prior to rotation (Thompson, 1989b), the first five eigenvalues of the correlation matrix were: 18.38346, 1.80270, 0.45648, 0.19292, 0.14571. This differed from the five factor solution which Newborg et al. (1984) derived from their BDI pilot norming data. Pre- and post-rotated factor results of this analysis are presented in Table 10 and 11.

Insert Tables 10 and 11 about here

Due to the high degree of similarity between the factor structures reported in Tables 7 and 9, we concluded that we had empirical support to derive reliability estimates using each test administration as a separate case. Results of this generalizability analyses indicated that internal reliability estimates for items contained within factors were high. Tables 12 through 14 provide results of these analyses. Factor I (98

items) yielded a generalizability coefficient of .98287, Factor II (51 items) yielded a generalizability coefficient of .96989, and Factor III (82 items) yielded a generalizability coefficient of .98300. We were unable to compare these results with the normative sample results because the developers of the BDI did not provide internal consistency reliability coefficients.

Insert Table 12 to 14 about here

Estimated model variance components in all three analyses were greater across items than persons. As Tables 12 through 14 indicate, model variance components for persons (P) were .172, .202, and .247 respectively. Item variance components (I) were .339, .356, and .376. This minimal variability among persons may reflect the homogeneity of the study sample. These results also suggest that children's "universe scores" vary only slightly from observed test scores.

Overall, results indicated high internal reliability for items contained in the subdomains which comprised each factor. We note, however, that the large G coefficients obtained in these analyses should be interpreted with a degree of caution. These coefficients should be further investigated in future research.

Discussion

Results of the measurement integrity analyses demonstrate that indices of item reliability can, and should, be obtained for items within factors. Generalizability analyses indicated high

internal consistency reliability estimates across the three obtained factors. Items comprising each of the subdomains across these factors were highly related to all other items on a particular factor.

The validity analysis disclosed differences in the underlying factor structure of the BDI. While the developers reported that the BDI contains five major domains, our analysis indicate that only three domains emerged when used with a sample of children with disabilities. This finding could have implications for researchers attempting to evaluate the effects of various interventions on the developmental status of young children with disabilities when using the BDI as a dependent measure. The developers of the BDI assert that each domain of the Battelle "stands alone" and may be administered independently. Our results indicate that for children similar to those involved in the present study, each of these original domains does not stand alone.

This finding supports the necessity for evaluating the construct validity of an instrument when the current sample and the normative sample differ on important characteristics. Failure to determine validity in such a situation could lead researchers and practitioners to assume that the test is measuring something that it is not.

This statement is especially relevant in the field of early childhood special education. As Sheehan (1982) explained:

> ... In most cases in which psychometric data are reported for infant assessment measures, the data reflect nonhandicapped infants. This factor becomes crucial when the instruments are used for purposes other than identification. It could be argued that data on nonhandicapped infants should form the basis for a comparison to determine if other infants are atypical. A problem occurs if, once identification is made, data on nonhandicapped children are used to compare progress with handicapped children. At present, this problem is not completely resolvable because appropriate instruments for atypical children do not exist and infant intervention projects are forced to use as a reference the normal standardization samples. (p. 49)

The measurement integrity analyses conducted in the present study further provide an example of the interesting relationship between validity and reliability. A valid test is always reliable, but a reliable test is not always valid. As Gay (1981) noted, "A reliable test can consistently measure the wrong thing and be invalid" (p.117).

In summary, we offer three implications for researchers in the field of early childhood special education. First, researchers should thoughtfully and explicitly compare the characteristics of proposed samples with the characteristics of those samples previously used to determine the integrity of

measurement tools. Second, researchers should usually conduct measurement integrity studies for the data in hand, rather than merely relying on an a priori premise that meaningful data were collected in substantive studies. Third, measurement integrity analyses are most critical when newly-developed instruments are used to operationalize study variables.

Throughout this paper we have emphasized that measurement integrity is not inherent within tests, but rather is a characteristic of the entire measurement process. This tenet should guide the conduct of all early childhood special education research. Our ability as researchers to detect clinically significant findings for the benefit of children with disabilities and their families is tied directly to the integrity of our measurement tools.

References

Brennan, R. (in press). Variance components in generalizability theory. In C.R. Reynolds (Ed.), Advances in cognitive assessment: An interdisciplinary perspective. New York: Plenum.

Boyd, R.D., Welge, P.A. Sexton, J.D., & Miller, J. (1989). Concurrent validity of the Battelle Developmental Inventory: Relationship with the Bayley Scales in young children with known or suspected disabilities. Journal of Early Intervention, 13, 14-23.

Crick, J.E., & Brennan, R.L. (1983). Manual for GENOVA: A generalized analysis of variance system (ACT Technical Bulletin No. 43). Iowa City, IA: The American College Testing Program.

Doll, E. (1965). Vineland Social Maturity Scale. Circle Pines, MN: American Guidance Services.

Gay, L.R. (1981). Educational research: Competencies for analysis and application. Columbus: Merrill.

Kerlinger, F.N. (1988). Foundations of behavioral research (3rd ed.). Fort Worth: Holt, Rinehart, & Winston.

Locke, L.F., Spirduso, W.W., & Silverman, S.J. (1987). Proposals that work: A guide for planning dissertations and grant proposals (2nd ed.). Newbury Park, CA: Sage.

Newborg, J., Stock, J., Wnek, L., Guidubaldi, J., & Svinicki, J. (1984). Battelle Developmental Inventory: Examiner's manual. Dallas, TX: DLM/Teaching Resources.

Nunnally, J. (1972). Educational measurement and evaluation (2nd ed.). New York: McGraw Hill.

Payton, O.D. (1988). Research: The validation of clinical practice. Philadelphia: F. A. Davis Company.

Sax, G. (1980). Principles of educational and psychological measurement and evaluation (2nd ed.). Belmont, CA: Wadsworth.

Sexton, D., Thompson, B., Perez, J., & Rheams, T. (1990). Maternal versus professional estimates of developmental status for young children with handicaps: An ecological approach. Topics in Early Childhood Special Education, 10 (3), 80-95.

Sexton, J.D., Thompson, B., Scott, R.L., & Wood, T.A. (1990). Measurement characteristics of the Inventory of Parent Experiences Scales. Topics in Early Childhood Education, 10(1), 36-49.

Shavelson, R.J., Webb, N.M., & Rowley, G.L. (1989). Generalizability theory. American Psychologist, 44, 922-932.

Sheehan, R. (1982). Infant assessment: A review and identification of emergent trends. In D.D. Bricker (Ed.), Intervention with at-risk and handicapped infants (pp. 47-62). Baltimore: University Park Press.

Thompson, B. (1989a, January). Why generalizability coefficients are an essential aspect of reliability assessment. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX.

Thompson, B. (1989b). Prerotation and postrotation eigenvalues shouldn't be confused: A reminder [editorial]. Measurement and Evaluation in Counseling and Development, 22(3), 114-116.

Thompson, B. (in press-a). Review of Generalizability theory: A primer by R.J. Shavelson & N.W. Webb. Educational and Psychological Measurement, 51.

Thompson, B. (in press-b). A partial test distribution for cosines among factors across samples. In B. Thompson (Ed.), Advances in social science methodology (Vol. 2). Greenwich, CT: JAI Press.

Table 1

Distribution of Norming Sample by Age, Race, and Sex and Mean Age of Each Age Group

Age Group[a]	Mean Age[a]	White		Minority		Total
		Male	Female	Male	Female	
0-5	3.0	20	22	3	5	50
6-11	8.4	26	19	2	2	49
12-17	14.5	22	21	1	5	49
18-23	20.3	25	22	5	2	54
24-35	29.9	39	43	9	8	99
36-47	41.6	43	39	6	14	102
48-59	54.2	41	44	8	8	101
60-71	65.7	38	45	8	7	98
72-83	77.0	42	44	13	9	108
84-95	89.0	34	42	8	6	90
TOTAL		330	341	63	66	800

Total by Race	671	129
BDI Percent	83.9	16.1
U.S. Percent	83.2	16.8

Overall Sex Distribution:	Male	Female
BDI Percent	49.1	50.9
U.S. Percent	48.6	51.4

[a]Age in months

Note. From Battelle Developmental Inventory examiner's manual (p. 51) by J. Newborg, J.R. Stock, L. Wnek, J. Guidubaldi, & J. Svinicki, 1984. Allen, TX:DLM. Copyright 1984 by LINC Associates.

Table 2

Distribution of Study Sample by Age, Race, and Sex and Mean Age of Each Age Group[a]

		White		Minority		
Age Group[b]	Mean Age[b]	Male	Female	Male	Female	Total
0-5	*	0	0	0	0	0
6-11	11.0	1	0	1	1	3
12-17	14.8	5	1	5	6	17
18-23	20.8	5	3	12	11	31
24-35	29.3	21	6	26	33	86
36-47	38.6	11	2	18	19	50
48-59	52.8	6	2	12	19	39
60-71	63.3	7	5	9	9	30
72-83	73.8	3	1	2	3	9
84-95	*	0	0	0	0	0
TOTAL		59	20	85	101	265

Total by Race	79	186
Study Sample Percent	29.9	70.1

Overall Sex Distribution:	Male	Female
Study Sample Percent	54.3	45.7

[a]Duplicated count across five years

[b]Age in months

Table 3

Correlations Among BDI Subdomain Scores

BDI Component	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Personal-Social Domain																						
1. Adult Interaction																						
2. Expressions/Feelings/Affect	.91																					
3. Self-Concept	.86	.92																				
4. Peer Interaction	.82	.90	.92																			
5. Coping	.73	.81	.84	.91																		
6. Social Role	.72	.83	.87	.92	.91																	
Adaptive Domain																						
7. Attention	.89	.86	.83	.79	.69	.69																
8. Eating	.93	.90	.87	.83	.73	.72	.90															
9. Dressing	.84	.90	.93	.91	.83	.87	.80	.86														
10. Personal Responsibility	.61	.71	.73	.83	.89	.89	.58	.62	.75													
11. Toileting	.76	.84	.92	.90	.84	.89	.75	.79	.93	.76												
Motor Domain																						
12. Muscle Control	.77	.66	.58	.52	.42	.40	.80	.81	.56	.33	.46											
13. Body Coordination	.79	.86	.88	.92	.91	.93	.76	.81	.88	.89	.87	.53										
14. Locomotion	.85	.90	.93	.92	.87	.90	.84	.88	.93	.80	.91	.62	.93									
15. Fine Muscle	.73	.82	.83	.89	.90	.92	.71	.75	.84	.92	.84	.45	.96	.89								
16. Perceptual Motor	.92	.90	.87	.82	.72	.71	.88	.96	.86	.61	.78	.78	.81	.87	.74							

(table continues)

BDI Component	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Communication Domain																						
17. Receptive	.76	.84	.87	.91	.92	.95	.74	.77	.86	.91	.88	.47	.95	.92	.95	.75						
18. Expressive	.80	.87	.90	.93	.90	.93	.78	.82	.89	.87	.89	.53	.94	.93	.93	.81	.96					
Cognitive Domain																						
19. Perceptual Discrimination	.85	.89	.92	.91	.86	.89	.85	.87	.91	.79	.90	.62	.92	.95	.88	.85	.91	.92				
20. Memory	.83	.87	.89	.90	.86	.88	.82	.86	.88	.81	.87	.62	.91	.93	.88	.84	.91	.92	.93			
21. Reasoning/Academic Skills	.67	.77	.79	.87	.90	.93	.65	.68	.80	.93	.81	.39	.93	.86	.95	.67	.95	.92	.85	.86		
22. Conceptual Development	.65	.75	.77	.86	.90	.93	.62	.66	.78	.93	.80	.35	.92	.84	.95	.64	.95	.91	.84	.85	.96	

Note. From Battelle Developmental Inventory examiner's manual (p. 61) by J. Newborg, J.R. Stock, L. Wnek, J. Guidubaldi, & J. Svinicki, 1984. Allen, TX:DLM. Copyright 1984 by LINC Associates.

Table 4

Correlations Among BDI Component Scores using Pooled Correlation Matrix across Five Years

BDI Component	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Personal-Social Domain																						
1. Adult Interaction																						
2. Expressions/Feelings/Affect	.77																					
3. Self-Concept	.69	.71																				
4. Peer Interaction	.65	.71	.77																			
5. Coping	.60	.65	.70	.71																		
6. Social Role	.32	.43	.67	.55	.46																	
Adaptive Domain																						
7. Attention	.74	.66	.58	.54	.53	.29																
8. Eating	.76	.63	.64	.60	.59	.34	.67															
9. Dressing	.57	.57	.65	.58	.55	.47	.47	.71														
10. Personal Responsibility	.45	.48	.65	.58	.58	.72	.44	.56	.56													
11. Toileting	.38	.40	.57	.48	.46	.46	.26	.43	.65	.48												
Motor Domain																						
12. Muscle Control	.66	.53	.41	.40	.40	.11	.61	.75	.61	.34	.34											
13. Body Coordination	.63	.59	.60	.58	.56	.37	.51	.77	.80	.54	.58	.77										
14. Locomotion	.63	.55	.48	.48	.52	.19	.47	.78	.72	.44	.50	.82	.91									
15. Fine Muscle	.70	.64	.72	.65	.62	.52	.64	.83	.76	.70	.54	.69	.79	.73								
16. Perceptual Motor	.67	.64	.71	.62	.61	.54	.60	.78	.72	.66	.52	.62	.79	.73	.83							

(table continues)

BDI Component	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Communication Domain																						
17. Receptive	.74	.69	.77	.72	.62	.65	.65	.70	.60	.65	.47	.50	.62	.51	.74	.74						
18. Expressive	.71	.68	.82	.70	.65	.67	.57	.71	.63	.70	.49	.47	.60	.50	.78	.73	.85					
Cognitive Domain																						
19. Perceptual Discrimination	.75	.66	.74	.62	.60	.52	.77	.75	.65	.64	.50	.66	.68	.60	.83	.79	.79	.76				
20. Memory	.76	.69	.77	.69	.60	.55	.74	.70	.60	.62	.46	.59	.66	.57	.78	.77	.79	.76	.83			
21. Reasoning/Academic Skills	.61	.59	.66	.62	.54	.65	.58	.63	.55	.66	.36	.45	.57	.46	.74	.74	.76	.74	.72	.74		
22. Conceptual Development	.52	.54	.70	.61	.53	.77	.47	.53	.55	.72	.46	.31	.51	.36	.65	.66	.80	.73	.69	.68	.74	

Note. Components obtained from the study sample are organized above into the five domains provided by the BDI developers for comparative purposes only.

Table 5

Correlations Between BDI Component Scores Obtained from Study Sample across Five Years

BDI Component	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Personal-Social Domain																						
1. Adult Interaction																						
2. Expressions/Feelings/Affect	.81																					
3. Self-Concept	.73	.78																				
4. Peer Interaction	.72	.79	.83																			
5. Coping	.68	.75	.78	.79																		
6. Social Role	.43	.54	.75	.64	.58																	
Adaptive Domain																						
7. Attention	.78	.72	.64	.62	.62	.40																
8. Eating	.80	.71	.69	.68	.68	.44	.71															
9. Dressing	.65	.67	.74	.68	.67	.58	.56	.75														
10. Personal Responsibility	.54	.60	.73	.68	.68	.77	.53	.63	.65													
11. Toileting	.49	.54	.69	.61	.60	.58	.38	.53	.74	.59												
Motor Domain																						
12. Muscle Control	.66	.53	.42	.42	.42	.16	.62	.74	.60	.36	.36											
13. Body Coordination	.69	.67	.68	.66	.65	.48	.58	.80	.83	.62	.65	.76										
14. Locomotion	.69	.65	.60	.59	.63	.34	.56	.82	.77	.55	.59	.80	.92									
15. Fine Muscle	.76	.72	.77	.73	.70	.61	.70	.85	.80	.75	.63	.69	.82	.78								
16. Perceptual Motor	.73	.72	.76	.70	.70	.62	.66	.81	.77	.72	.61	.62	.82	.78	.86							

(table continues)

BDI Component	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Communication Domain																						
17. Receptive	.78	.76	.82	.78	.71	.72	.70	.74	.69	.72	.59	.51	.69	.60	.79	.79						
18. Expressive	.76	.76	.86	.78	.74	.74	.64	.76	.72	.76	.62	.48	.68	.61	.82	.79	.88					
Cognitive Domain																						
19. Perceptual Discrimination	.79	.74	.79	.71	.69	.60	.80	.79	.72	.70	.61	.65	.74	.68	.86	.83	.83	.81				
20. Memory	.79	.75	.82	.75	.69	.64	.78	.74	.68	.69	.58	.58	.72	.65	.82	.81	.83	.81	.86			
21. Reasoning/Academic Skills	.66	.66	.73	.69	.63	.72	.64	.67	.64	.72	.49	.46	.64	.55	.78	.78	.81	.79	.77	.79		
22. Conceptual Development	.60	.65	.78	.70	.65	.82	.55	.59	.66	.77	.60	.33	.60	.49	.71	.72	.85	.80	.75	.75	.75	

Note. Components obtained from the study sample are organized above into the five domains provided by the BDI developers for comparative purposes only.

Table 6

Pre-rotated Factor Structure Matrix, using Intercorrelations among Subscales with All Cases Treated Independently

	FACTOR I	FACTOR II	FACTOR III	h^2
FM	.92344	.11232	.06598	.86971
PD	.91010	.05245	-.13929	.85044
EX	.90575	-.19093	-.06975	.86171
PM	.90526	.07681	.08253	.83220
R	.90393	-.16841	-.14991	.86793
MEM	.89996	-.02728	-.18385	.84447
SC	.89407	-.23335	-.00672	.85386
E	.86685	.31247	-.02365	.84963
BC	.85263	.31955	.29055	.91351
PI	.84793	-.15632	-.05593	.74655
AI	.84620	.21714	-.29491	.85017
EFA	.84535	.03165	-.19393	.75323
D	.84502	.10367	.34321	.84259
RA	.84035	-.19885	-.12834	.76220
CD	.82635	-.39471	.00793	.83871
C	.81917	-.09349	-.00187	.67978
PR	.80339	-.30123	.13355	.75401
L	.79316	.46721	.27566	.92338
A	.77185	.18962	-.44737	.83186
SR	.71578	-.58172	.12469	.86629
T	.70745	-.12034	.48681	.75195
MC	.65884	.64497	-.00988	.85016

Table 7

Post-rotated Factor Structure Matrix, using Intercorrelations among Subscales with all Cases Treated Independently

	FACTOR I	FACTOR II	FACTOR III	h^2
SR	.91511	.14733	.08463	.86629
CD	.82425	.34935	.19308	.83871
PR	.76718	.26140	.31164	.75401
SC	.74615	.44020	.32148	.85386
EX	.70921	.50752	.31804	.86171
R	.67459	.57578	.28519	.86793
RA	.66056	.51561	.24496	.76220
PI	.64942	.47351	.31717	.74655
T	.64676	-.02451	.57710	.75195
C	.59607	.43105	.37238	.67978
A	.26362	.83063	.26911	.83186
AI	.32338	.75827	.41306	.85017
MEM	.56112	.63694	.35202	.84447
EFA	.48049	.62991	.35436	.75323
PD	.51842	.62792	.43289	.85044
L	.22550	.34205	.86921	.92338
BC	.37552	.32501	.81662	.91351
MC	-.05186	.53960	.74586	.85016
D	.54044	.22328	.70758	.84259
E	.32373	.57970	.63935	.84963
FM	.52622	.48834	.59525	.86971
PM	.54410	.45599	.57291	.83220
Trace	15.46668	1.74229	.98536	18.19433

Table 8

Pre-rotated Factor Structure Matrix, using Interrcorrelations among Subscales from the Pooled Correlation Matrix

	FACTOR I	FACTOR II	FACTOR III	h^2
FM	.90824	-.11090	.08776	.84490
PD	.88965	-.03285	-.13838	.81171
PM	.88517	-.07270	.10734	.80033
R	.88063	.20697	-.14390	.83905
MEM	.87857	.03854	-.19857	.81281
EX	.87245	.22560	-.06306	.81603
SC	.85760	.24241	-.01660	.79451
E	.85133	-.31116	-.02647	.82228
BC	.82248	-.36642	.29343	.89684
AI	.81995	-.20080	-.33883	.82745
RA	.80810	.23227	-.10676	.71837
D	.79287	-.17292	.37340	.79797
EFA	.79013	-.03768	-.26253	.69464
PI	.78957	.17326	-.07932	.65973
CD	.77126	.43973	.04332	.79007
PR	.74960	.34071	.20259	.71902
C	.74555	.07887	-.03173	.56308
L	.73728	-.54890	.25595	.91038
A	.73384	-.16366	-.47794	.79373
MC	.68688	-.60522	-.02935	.83896
SR	.63715	.63608	.19818	.84983
T	.60499	.04735	.54948	.67019

Table 9

Post-rotated Factor Structure Matrix, using Interrcorrelations among Subscales from the Pooled Correlation Matrix

	FACTOR I	FACTOR II	FACTOR III	h^2
SR	.91529	.09142	.06094	.84983
CD	.80770	.33322	.16328	.79007
PR	.76030	.21836	.30542	.71902
SC	.69487	.47332	.29603	.79451
EX	.67922	.52170	.28728	.81603
R	.64940	.59305	.25615	.83905
RA	.63445	.51560	.22362	.71837
PI	.58561	.49595	.26613	.65973
C	.49980	.45306	.32866	.56308
A	.19528	.84281	.21279	.79373
AI	.25438	.79468	.36223	.82745
EFA	.37946	.68300	.29011	.69464
MEM	.50667	.67011	.32720	.81281
PD	.47461	.64547	.41209	.81171
L	.09272	.36007	.87871	.91038
BC	.29170	.34296	.83314	.89684
D	.44086	.22207	.74451	.79797
MC	-.05298	.56251	.72093	.83896
E	.26932	.59562	.62848	.82228
T	.53966	-.07337	.61121	.67019
FM	.48433	.49850	.60151	.84490
PM	.50439	.46141	.57708	.80033
TRACE	14.08533	1.97621	1.21034	

Table 10

Pre-rotated Factor Structure Matrix, using Interrcorrelations among Subscales as Reported in Table 3

	FACTOR I	FACTOR II	h^2
L	.97112	.04355	.94498
EX	.96674	-.11367	.94752
BC	.96418	-.13755	.94857
PD	.96301	.04847	.92973
PI	.95826	-.05849	.92169
R	.95678	-.21705	.96253
MEM	.95337	.02177	.90939
SC	.94404	.10073	.90136
D	.93884	.06758	.88598
FM	.93681	-.24400	.93716
SR	.93559	-.26503	.94557
EFA	.93146	.19557	.90587
T	.91981	-.05664	.84925
C	.91766	-.23318	.89647
RA	.90830	-.33751	.93892
CD	.89493	-.37786	.94367
E	.89353	.39363	.95334
PM	.88244	.39096	.93155
AI	.87741	.37971	.91403
PR	.85968	-.40054	.89948
A	.85361	.40433	.89214
MC	.61988	.66537	.82697

Table 11

Post-rotated Factor Structure Matrix, using Intercorrelations among Subscales as Reported in Table 3

	FACTOR I	FACTOR II	h^2
CD	.93521	.26279	.94367
PR	.92180	.22307	.89948
RA	.92049	.30267	.93892
SR	.89662	.37635	.94557
FM	.88446	.39355	.93716
R	.88326	.42706	.96253
C	.86274	.39006	.89647
BC	.83947	.49383	.94857
EX	.82658	.51409	.94752
PI	.78553	.55193	.92169
T	.75432	.52939	.84925
L	.73194	.63972	.94498
MEM	.73165	.61162	.90939
PD	.72253	.63849	.92973
D	.69172	.63836	.88598
SC	.67511	.66752	.90136
MC	.06957	.90671	.82697
E	.45296	.86497	.95334
PM	.44594	.85597	.93155
A	.41507	.84844	.89214
AI	.44903	.84404	.91403
EFA	.60613	.73381	.90587
TRACE	18.38346	1.80270	

Table 12

Generalizability Results for Factor I - 98 Items

EFFECT	DEGREES OF FREEDOM	SUMS OF SQUARES FOR MEAN SCORES	SUMS OF SQUARES FOR SCORE EFFECTS	MEAN SQUARES	(QF = QUASI F RATIO) F STATISTIC	F-TEST DEGREES OF FREEDOM NUMERATOR	DENOMINATOR
P	264	14166.84694	4531.08410	17.16320	58.38327	264	25608
I	97	18371.81509	8736.05225	90.06239	306.36117	97	25608
PI	25608	30431.00000	7528.10081	.29397			
TOTAL	25969		20795.23716				

Note. For Generalizability analyses, F-statistics should be ignored.

MODEL VARIANCE COMPONENTS

EFFECT	DEGREES OF FREEDOM	USING ALGORITHM	USING EMS EQUATIONS	STANDARD ERROR
P	264	.1721349	.1721349	.0151861
I	97	.3387488	.3387488	.0483053
PI	25608	.2939746	.2939746	.0025979

Note. The Algorithm & EMS estimated variance components will be identical if no negative estimates.

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	.17213	.41489	1.00000	
EXPECTED OBSERVED SCORE	.17513	.41849	1.00005	
LOWER CASE DELTA	.00300	.05477	.01020	GENERALIZABILITY COEFFICIENT = .98287 (57.38327)
UPPER CASE DELTA	.00646	.08035	.01443	PHI = .96385 (26.66129)
MEAN	.00412	.06417		

Table 13

Generalizability Results for Factor II - 51 Items

EFFECT	DEGREES OF FREEDOM	SUMS OF SQUARES FOR MEAN SCORES	SUMS OF SQUARES FOR SCORE EFFECTS	MEAN SQUARES	(QF = QUASI F RATIO) F STATISTIC	F-TEST DEGREES OF FREEDOM NUMERATOR	DENOMINATOR
P	264	25040.50980	2803.22560	10.61828	33.20985	264	13200
I	50	26972.30189	4735.01768	94.70035	296.18595	50	13200
PI	13200	33996.00000	4220.47251	.31973			
TOTAL	13514		11758.71580				

Note. For Generalizability analyses, F-statistics should be ignored.

MODEL VARIANCE COMPONENTS

EFFECT	DEGREES OF FREEDOM	USING ALGORITHM	USING EMS EQUATIONS	STANDARD ERROR
P	264	.2019323	.2019323	.0180535
I	50	.3561533	.3561533	.0700840
PI	13200	.3197328	.3197328	.0039353

Note. The Algorithm & EMS estimated variance components will be identical if no negative estimates.

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	.20193	.44937	1.00000	
EXPECTED OBSERVED SCORE	.20820	.45629	1.00019	
LOWER CASE DELTA	.00627	.07918	.01961	GENERALIZABILITY COEFFICIENT = .96989 (32.20986)
UPPER CASE DELTA	.01325	.11512	.02773	PHI = .93841 (15.23710)
MEAN	.00777	.08814		

Table 14

Generalizability Results for Factor III - 82 Items

EFFECT	DEGREES OF FREEDOM	SUMS OF SQUARES FOR MEAN SCORES	SUMS OF SQUARES FOR SCORE EFFECTS	MEAN SQUARES	(QF = QUASI F RATIO) F STATISTIC	F-TEST DEGREES OF FREEDOM NUMERATOR	DENOMINATOR
P	264	23910.60976	5429.17395	20.56505	58.83279	264	21384
I	81	26572.03019	8090.59439	99.88388	285.74920	81	21384
PI	21384	39476.00000	7474.79586	.34955			
TOTAL	21729		20994.56420				

Note. For Generalizability analyses, F-statistics should be ignored.

MODEL VARIANCE COMPONENTS

EFFECT	DEGREES OF FREEDOM	USING ALGORITHM	USING EMS EQUATIONS	STANDARD ERROR
P	264	.2465305	.2465305	.0217466
I	81	.3756012	.3756012	.0585094
PI	21384	.3495509	.3495509	.0033803

Note. The Algorithm & EMS estimated variance components will be identical if no negative estimates.

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	.24653	.49652	1.00000	
EXPECTED OBSERVED SCORE	.25079	.50079	1.00007	
LOWER CASE DELTA	.00426	.06529	.01220	GENERALIZABILITY COEFFICIENT = .98300 (57.83279)
UPPER CASE DELTA	.00884	.09404	.01725	PHI = .96537 (27.87760)
MEAN	.00553	.07434		

Appendix

BDI Subdomain Key

1.	AI	Adult Interaction
2.	EFA	Expression of Feelings and Affect
3.	SC	Self Concept
4.	PI	Peer Interaction
5.	C	Coping
6.	SR	Social Role
7.	A	Attention
8.	E	Eating
9.	D	Dressing
10.	PR	Personal Responsibility
11.	T	Toileting
12.	MC	Muscle Control
13.	BC	Body Coordination
14.	L	Locomotion
15.	FM	Fine Muscle
16.	PM	Perceptual Motor
17.	R	Receptive
18.	EX	Expressive
19.	PD	Perceptual Discrimination
20.	MEM	Memory
21.	RA	Reasoning and Academic Skills
22.	CD	Conceptual Development

Authors' Notes

Some of these results were also presented at the Fourth Annual Allied Health Research Symposium, Birmingham, AL, October, 1991.